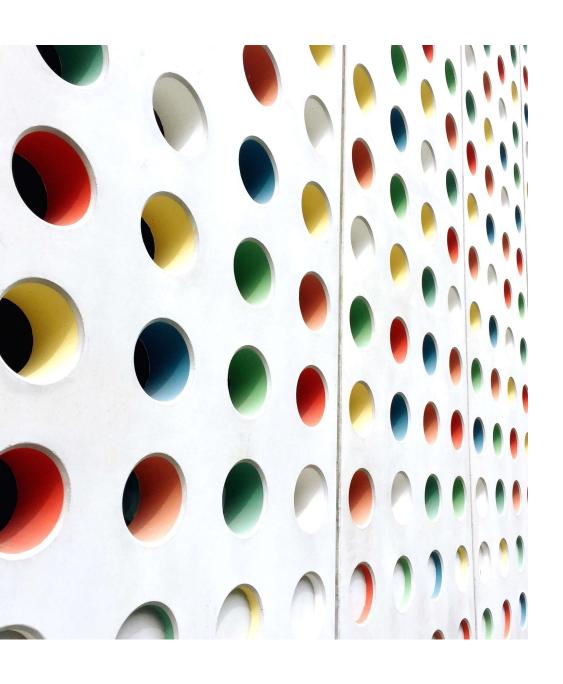
# Peranan ANALISIS FAKTOR dalam Analisis Butir

Wahyu Widhiarso | Fakultas Psikologi UGM



# Peranan Analisis Faktor dalam Seleksi Butir

# Tiga Analisis Utama dalam Psikometrika



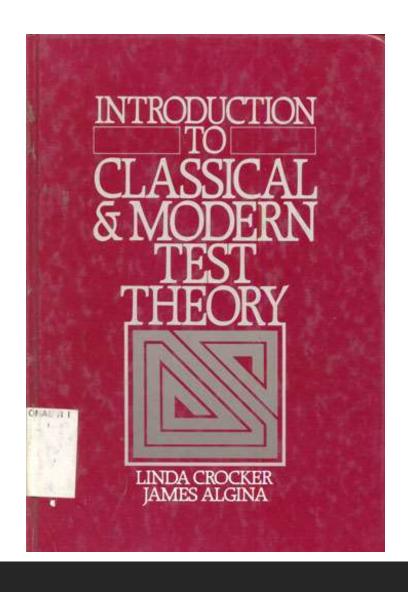
TEORI TES KLASIK



ANALISIS FAKTOR



TEORI RESPONS BUTIR



### UNIT IV Item Analysis in Test Development

### Chapter 14 Item Analysis 311

Item Difficulty, Mean, and Variance 311
Item Discrimination 313
Item Reliability and Validity Indices 320
Conducting an Item Analysis Study 321
Item Analysis for Criterion-Referenced Tests 329
Summary 335
Exercises 336

### Chapter 15 Introduction to Item Response Theory 339

Basic Concepts of Item Response Theory 340 The Normal Ogive 346 Relating Item Response Theory to Classical Test Theory 350 Logistic Models 352 Estimation of Parameters 355 Selection of a Model 361 362 Applications of Item Response Theory Summary 371 371 Exercises

### Teori Tes Klasik

- Biasa dipakai sebagai landasan untuk melakukan analisis butir
- Memiliki beberapa properti psikometris (daya beda butir, tingkat kesulitan butir, efektivitas distraktor, reliabilitas pengukuran dsb)

### Teori Tes Modern

- Biasa dipakai sebagai landasan untuk melakukan analisis butir
- Memiliki beberapa properti psikometris (daya beda butir, tingkat kesulitan butir, fungsi informasi butir-tes)

Teori Tes Klasik

Teori Tes Modern

- Scale Refinement
- Purification

- Item Calibration
- Scaling Linking Equating

# **Muncullah Analisis Faktor**



- Scale Refinement
- Purification

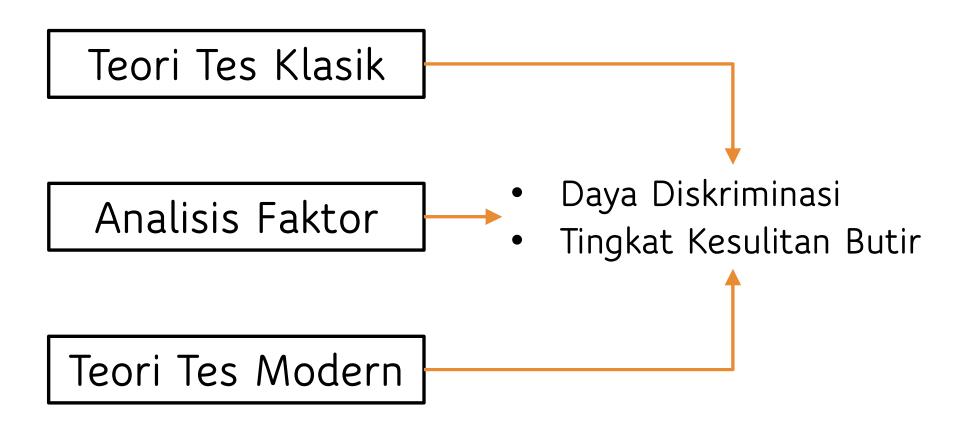
Analisis Faktor

Dimensionalitas

Teori Tes Modern

- Item Calibration
- Scaling Linking Equating

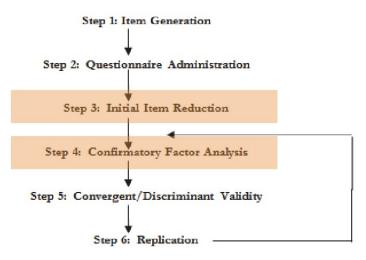
# Ketiganya memiliki Properti yang mirip



# Pendekatan 1

# Analisis Butir Tes Klasik

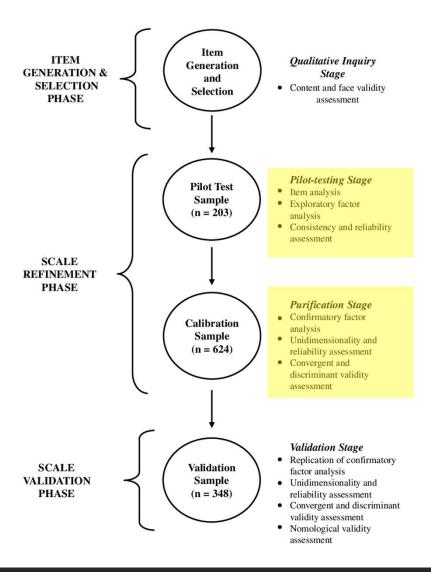
# Analisis Faktor

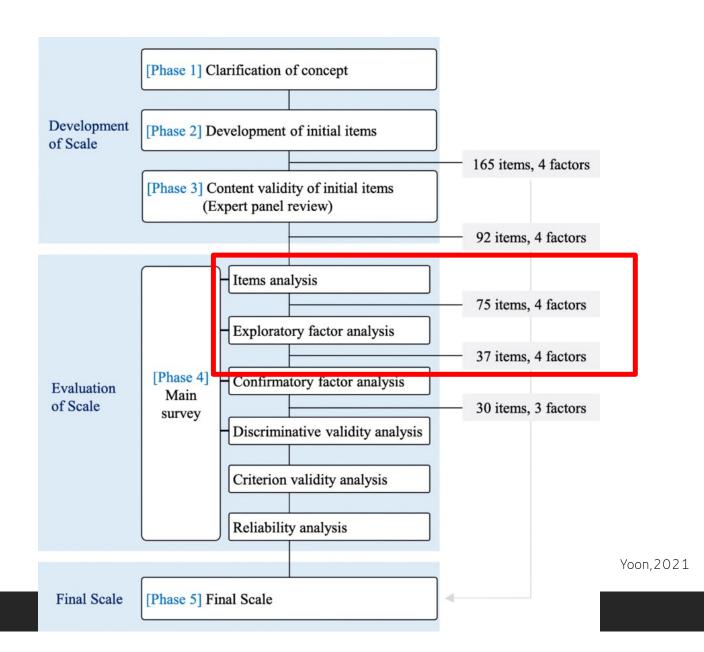


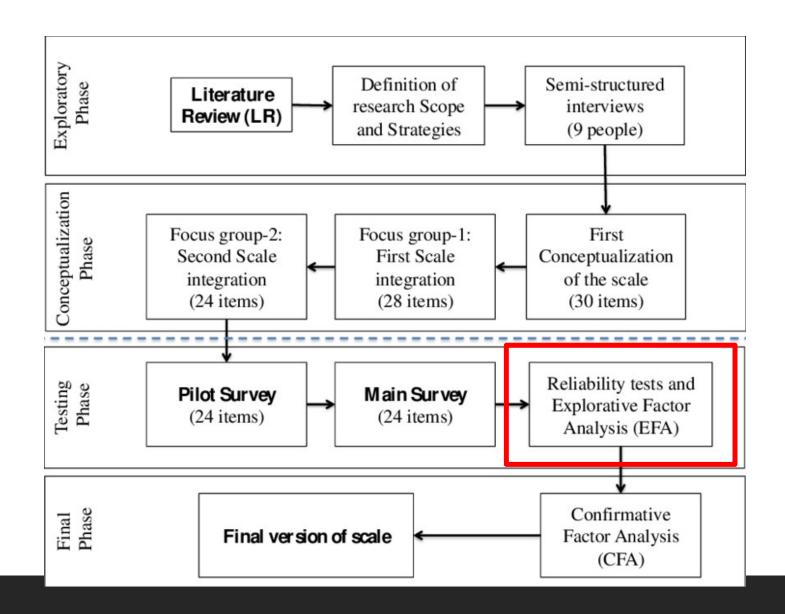
Hinkin, 1998

# Apa alasannya?

Analisis butir dulu baru pengujian validitas struktural







# Pendekatan 2

# Analisis Butir Tes Klasik

## Analisis Faktor

### Step 3: Designing and Conducting Studies to Develop and Refine the Scale Issues to Consider:

- (a) Pilot testing as an item-trimming procedure
- (b) The use of several samples from relevant populations for scale development
- (c) Designing the studies to test psychometric properties
- (d) Initial item analyses via exploratory factor analyses (EFAs)
- (e) Initial item analyses and internal consistency estimates
- (f) Initial estimates of validity
- (g) Retaining items for the next set of studies

Netemeyer, 2003

# Apa alasannya?

Analisis butir perlu dilakukan per dimensi sehingga analisis faktor diperlukan di awal

### Stage 1 Scale Generation and Initial Purification

Stage 2 Scale
Purification through
Exploratory Factor
Analysis (EFA)

### Stage 3 Reliability and Validity Assessment

Stage 4 Introduction EXQ Scale

- Using insight from literature review to choose context
- In-depth interviews (n=30)
- Generate initial pool of items
- -Readability Check (n=10)
- Expert judgment tasks to assess face and construct validity (n=7)
- -Q-sorting (8 categories)
- Initial purification and categorization (a)n=7
- (b)n=5
- Produced 37 items in 5 dimensions for next stage

Qualitative Enquiry and Initial Purification

- Collect responses from representative sample (n=75)
- Exploratory factor analysis (on 5-dimensional model)
- Assess content validity, scale reliability and validity
- Develop purified scale with 19 items in four dimensions for next stage

Purification and Refinement

- Collect responses from additional representative sample (n=218)
- Confirmatory factor analysis (on 4-dimensional model)
- Assess factor structure and dimensionality of scale
- Assess model fit
   Assess scale and construct reliability and discriminant validity
- Final EXQ scale of 19 items in four dimensions

 Final EXQ scale of 19 items in four dimensions, conceptual model

Final Refinement and Validation

# Pendekatan 3

# Analisis Faktor

# Apa alasannya?

Analisis faktor memiliki fungsi sama dengan analisis butir namun lebih tepat karena sesuai dengan model

### Constructing Validity: Basic Issues in Objective Scale Development

### Lee Anna Clark and David Watson The University of Iowa

A primary goal of scale development is to create a valid measure of an underlying construct. We discuss theoretical principles, practical issues, and pragmatic decisions to help developers maximize the construct validity of scales and subscales. First, it is essential to begin with a clear conceptualization of the target construct. Moreover, the content of the initial item pool should be overinclusive and item wording needs careful attention. Next, the item pool should be tested, along with variables that assess closely related constructs, on a heterogeneous sample representing the entire range of the target population. Finally, in selecting scale items, the goal is unidimensionality rather than internal consistency; this means that virtually all interitem correlations should be moderate in magnitude. Factor analysis can play a crucial role in ensuring the unidimensionality and discriminant validity of scales.

Scale development remains a growth industry within psychology. A PsycLIT database survey of articles published in the 6-year period from 1989 through 1994 revealed 1,726 articles with the key words "test construction" or "scale development" published in English-language journals, 270 in other-language journals, and 552 doctoral dissertations. During this same period (i.e., beginning with its inception), 50 articles addressing scale development or test construction were published in *Psychological Assessment* alone. The majority of these articles reported the development of one or more new measures (82%); most of the rest presented new scales derived from an existing instrument (10%). We use these 41 scale-development articles

signed to be completed by clinicians, parents, teachers, spouses, peers, and so forth.

Before proceeding further, it is interesting to examine the new measures comprising our *Psychological Assessment* sample. This examination sample offers a glimpse at why scale development continues unabated, as well as the nature of the unmet needs these scale developers are seeking to fill. First, not surprisingly given this journal's focus, more than half (61%) of the scales assess some aspect of psychopathology, personality, or adjustment. The next most common categories are measures of attitudes and interpersonal relations (20% and 15%, respectively). The remaining scales assess a miscellany of behaviors, abilities, response validity,



### Personality and Individual Differences Volume 12, Issue 3, 1991, Pages 291-294



Does item homogeneity indicate internal consistency or item redundancy in psychometric scales?

Gregory J. Boyle

Show more ✓

+ Add to Mendeley ≪ Share 55 Cite

https://doi.org/10.1016/0191-8869(91)90115-R Get rights and content

### Abstract

The term 'internal consistency' has been used extensively in classical psychometrics to refer to the reliability of a scale based on the degree of within-scale item intercorrelation, as measured by say the split-half method, or more adequately by Cronbach's (1951) (*Psychometrika*, 16, 297–334) alpha, as well as the KR<sub>20</sub> and KR<sub>21</sub> coefficients. This term is a misnomer, as a high estimate of internal item consistency/item homogeneity may also suggest a high level of item redundancy, wherein essentially the same item is rephrased in several different ways.



### Specifying the domain

OSS projects hosted on Sourceforge



### Generating of initial scale items

Literature review, Studying Sourceforge.net, Examining possible queries on enhancement process; Result=9 initial items



### Data Collection: first stage

Collecting data on the 9 items from 240 OSS projects which adopted enhancement tracking system; Result=data collected on 9 items from 240 projects



### Scale Purification



Assess convergent validity of items (CFA, Cronbach's alpha, AVE); Assess discriminant validity (CFA); Result= Convergent and discriminant validity were found for 6 out of 9 measures; list of 6 candidate measures



### Data Collection: second stage

Collecting data on the 6 items from 750 OSS projects to evaluate purified items; Result=data collected on the 6 items from 750 projects

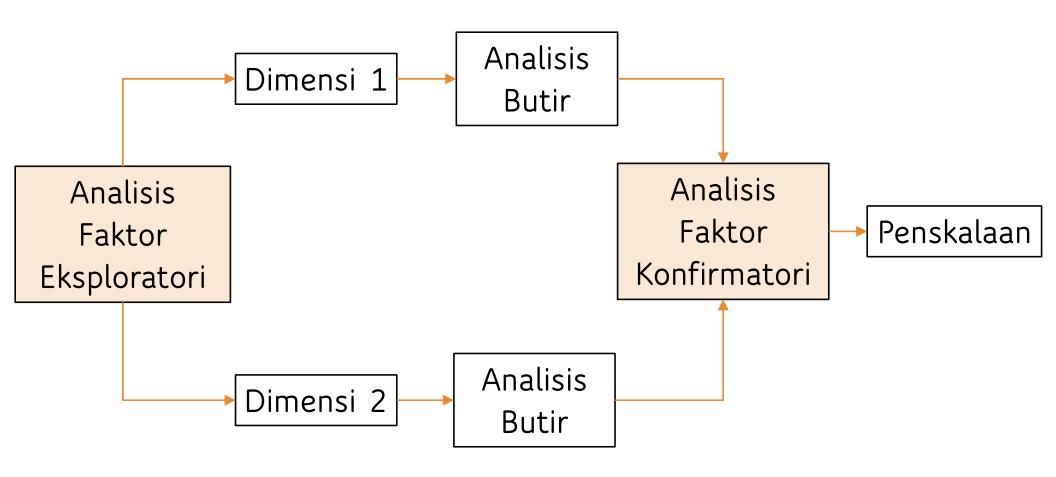


### Assessing validity and reliability



Assess composite reliabilities and AVE using PLS; Assess convergent and discriminant validity using PLS; Result = Convergent and discriminant validity were found for 4 of the purified measures; final list of 4 measures

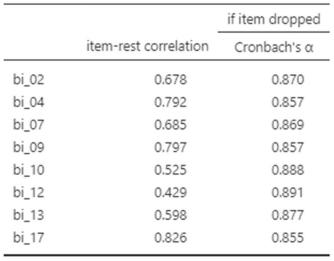
# Manakah yang perlu diikuti?

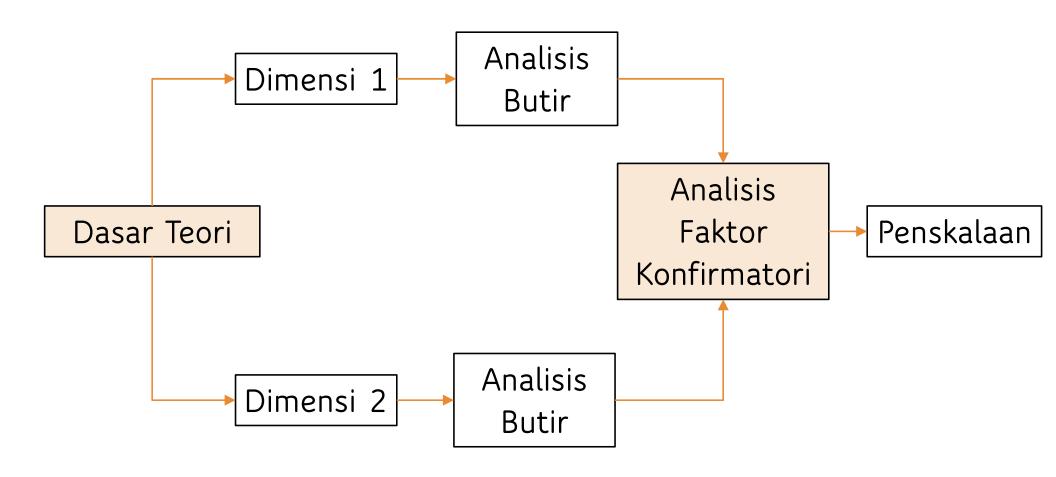


### Item Reliability Statistics

		if item dropped
	item-rest correlation	Cronbach's α
bi_01	0.7826	0.904
bi_02	0.0501	0.919
bi_03	0.7762	0.905
bi_04	0.1813	0.918
bi_05	0.7530	0.905
bi_06	0.8083	0.904
bi_07	0.0688	0.919
bi_08	0.8547	0.902
bi_09	0.1978	0.917
bi_10	0.0519	0.918
bi_11	0.7604	0.905
bi_12	0.1602	0.917
bi_13	0.0736	0.918
bi_14	0.8404	0.903
bi_15	0.8055	0.904
bi_16	0.8029	0.904
bi_17	0.1984	0.917
bi_18	0.7000	0.907
bi_19	0.8121	0.904









- Scale Refinement
- Purification

Analisis Faktor

Dimensionalitas

Teori Tes Modern

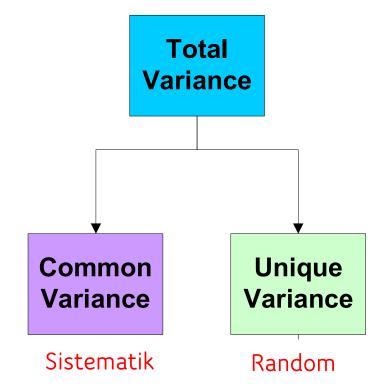
- Item Calibration
- Scaling Linking Equating

### Analisis Faktor

- Dipakai untuk mengidentifikasi struktur pengukuran
- Struktur pengukuran menunjukkan berapa faktor atau dimensi dalam suatu konstruk ukur alat atau alat ukur serta keterkaitan di antara dimensi tersebut
- Melalui analisis faktor peneliti mengembangkan model tertentu untuk memaksimalkan varians skor hasil pengukuran

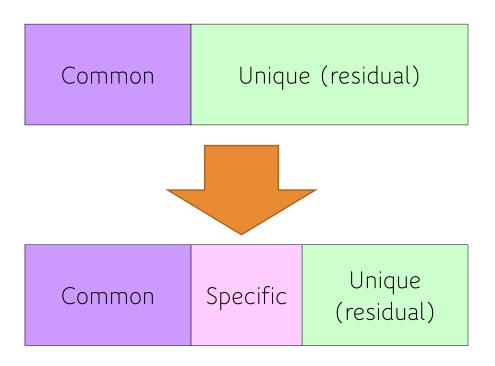
Total Variance

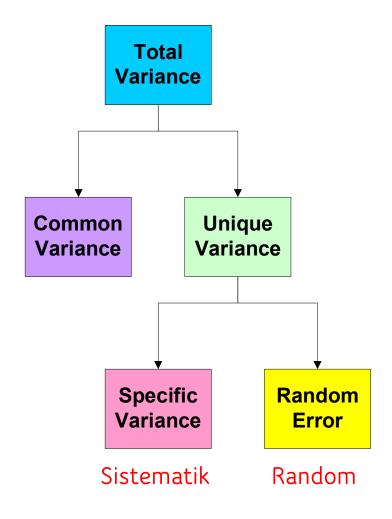
Common Unique (residual)



- 1. Saya suka berteman [Ya]
- 2. Saya suka menyapa orang lain [Ya]
- 3. Saya suka menolong [Tidak]

### Total Variance





# Ilustrasi

### Item Reliability Statistics

		if item dropped
	item-rest correlation	Cronbach's α
bi_01	0.7826	0.904
bi_02	0.0501	0.919
bi_03	0.7762	0.905
bi_04	0.1813	0.918
bi_05	0.7530	0.905
bi_06	0.8083	0.904
bi_07	0.0688	0.919
bi_08	0.8547	0.902
bi_09	0.1978	0.917
bi_10	0.0519	0.918
bi_11	0.7604	0.905
bi_12	0.1602	0.917
bi_13	0.0736	0.918
bi_14	0.8404	0.903
bi_15	0.8055	0.904
bi_16	0.8029	0.904
bi_17	0.1984	0.917
bi_18	0.7000	0.907
bi_19	0.8121	0.904

### Total Variance

Common Unique (residual)

### Item Reliability Statistics

		if item dropped
	item-rest correlation	Cronbach's α
bi_02	0.678	0.870
bi_04	0.792	0.857
bi_07	0.685	0.869
bi_09	0.797	0.857
bi_10	0.525	0.888
bi_12	0.429	0.891
bi_13	0.598	0.877
bi_17	0.826	0.855

### Total Variance

Com	Specific	Unique (residual)
		,

# Kapan Menggunakan Analisis Faktor?

- Mengidentifikasi struktur data
- Menguji validitas faktorial/konstruk/struktural
- Menguji validitas divergen/konvergen

# Semoga Bermanfaat